

Cuadernos del CIMBAGE N° 15 (2012) 19 - 31

UN ANÁLISIS DEL APOORTE DE ESTIMADORES GMM EN SISTEMAS Y DIFERENCIAS A LA ESTIMACIÓN DE ECUACIONES DINÁMICAS EN ECONOMÍA

Cintia Martínez

Sección de Investigaciones en Métodos Cuantitativos para la Gestión - IADCOM
Facultad de Ciencias Económicas, Universidad de Buenos Aires
Av. Córdoba 2122 – 2° piso – Ciudad de Buenos Aires – C1120AAQ - Argentina
cintiam34@yahoo.com

Recibido 5 de diciembre de 2012, aceptado 6 de febrero de 2013

Resumen

El estudio de la convergencia económica entre grupos de países o regiones ha sido testado económicamente en las últimas décadas con datos de panel, y en particular con modelos de efectos fijos y métodos de momentos generalizados en diferencias y en sistemas. El problema económico ha contribuido positivamente al desarrollo de la metodología de paneles dinámicos, pero el aporte que la metodología econométrica ha hecho a la comprobación empírica de la existencia de convergencia no resulta tan obvio. En este trabajo, se realiza una revisión del aporte de los métodos GMM en diferencias y en sistemas a la estimación econométrica de ecuaciones dinámicas y se concluye acerca de la pertinencia de este método para la modelización de convergencia económica.

Palabras clave: modelos de datos de panel, estimación con variables instrumentales, paneles dinámicos, system GMM, GMM-diff.

**AN ANALYSIS OF THE CONTRIBUTION OF GMM SYSTEM
AND GMM DIFFERENCES ESTIMATORS TO THE
ESTIMATION OF ECONOMIC DYNAMIC EQUATIONS**

Cintia Martínez

Sección de Investigaciones en Métodos Cuantitativos para la Gestión - IADCOM
Facultad de Ciencias Económicas, Universidad de Buenos Aires
Av. Córdoba 2122 – 2° piso – Ciudad de Buenos Aires – C1120AAQ - Argentina
cintiam34@yahoo.com

Received December 5th 2012, accepted February 6th 2013

Abstract

The study of economic convergence between countries or regions has been tested econometrically along the last decades with panel data models, in particular with fixed effects models and generalized method of moments in differences and systems. The economic problem has contributed in a positive way to the development of the econometric methodology, but the inverse is not that obvious. In this paper, the contribution of GMM differences and GMM system to the estimation of dynamic equations is revised and the pertinence of them to the modeling of economic convergence is discussed.

Keywords: models with panel data, instrumental variables estimation, dynamic panel data models, system GMM, GMM-diff.

1. INTRODUCCIÓN Y MOTIVACIÓN

Los estudios de crecimiento económico centraron su análisis en la posible existencia de un proceso de convergencia entre distintos países, estén agrupados en regiones o no. La existencia de convergencia beta ha sido testeada utilizando diversas técnicas estadísticas y econométricas. Estos estudios involucran automáticamente la evolución en el tiempo de variables económicas y de las ecuaciones que las relacionan.

Hacia fines de los 80, la investigación aplicada intentó confirmar la existencia o no de convergencia beta entre grupos de países o regiones a través de la utilización de modelos econométricos tipo cross-section. Son pioneros los estudios de Landau (1983), Kormendi y Meguire (1985), Barro (1989) y, con posterioridad, Barro y Sala-i-Martin (1995). Ya en los 90, se popularizó el uso de datos de panel y surgió el primer inconveniente importante: la existencia de errores correlacionados, introducidos por el dinamismo de la ecuación proveniente de la teoría económica.

Bhargava y Sargan (1983 y 1986) puntualizaron los problemas derivados de estimar una ecuación dinámica en datos de panel y propusieron el uso de un sistema de ecuaciones simultáneas con estimación máximo merosímiles. Posteriormente, propusieron para paneles dinámicos la construcción de una matriz de variables instrumentales con los valores pasados, presentes y futuros de las variables estrictamente exógenas. En particular, su metodología se aplicó a paneles de dimensión temporal corta.

Anderson y Hsiao (1981) estudiaron paneles en los cuales ambas dimensiones tienden a infinito, con estimadores MLE. Estos autores y Griliches y Hausman (1986) postularon como supuesto la existencia de mínima o limitada correlación entre las variables rezagadas de la independiente, las cuales fueron utilizadas como instrumentos.

Arellano y Bond (1991) propusieron estimadores generalizados de momentos (GMM) para datos de panel con efectos fijos, y desarrollaron tests para contrastar la existencia de autocorrelación de primero y segundo orden, basados en Sargan y Hausman.

Arellano-Bover (1995) y Blundell-Bond (1998) abordaron el problema de la presencia de series múltiples no estacionarias en datos de panel, aumentando la matriz de variables instrumentales. Desarrollaron el método GMM-System para el caso en que las observaciones muestren un comportamiento cercano al de random walk.

Este trabajo se estructura de la siguiente manera: se especifica brevemente el análisis de convergencia económica derivado del modelo basal de Solow y Swan (1956) y la pregunta fundamental acerca de la existencia o no de convergencia condicional; la estimación econométrica de esta convergencia requiere la consideración de la dinámica en el tiempo de su ecuación básica, por lo que se enumeran los problemas derivados de esta dinámica y las soluciones que la literatura ha propuesto. A continuación, se discuten las soluciones aportadas por la literatura y se concluye acerca del aporte de esta metodología econométrica a la pregunta básica acerca de existencia o no de convergencia condicional, aplicando los estimadores estudiados al caso de los países del Mercosur durante 1991-2008.

2. LA CONVERGENCIA ECONÓMICA EN LOS MODELOS DE CRECIMIENTO ENDÓGENO

La pregunta que motiva fundamentalmente el análisis de convergencia económica es: *¿pueden los países pobres alcanzar a los países ricos en el largo plazo?*

Con este objetivo, se definen:

Convergencia sigma: la dispersión alrededor de la media del PBI per cápita de todos los países o regiones considerados en el estudio. Se puede medir a través del desvío estándar o del coeficiente de variación.

Convergencia beta: si los países o regiones con PBI p. c. inicial más bajo crecen a tasas mayores que los demás, convergerán con el resto en un plazo determinado. La convergencia se da cuando existe una relación inversa entre el valor inicial del PBI p.c. y su tasa de crecimiento.

Convergencia beta condicional: los países pobres alcanzarán a los ricos siempre y cuando reúnan ciertas condiciones al inicio que les permitan hacerlo.

En este estudio trabajaremos sobre la base del tercer concepto de convergencia.

3. ESTIMACIÓN CON DATOS DE PANEL

Derivado del modelo neoclásico de Solow y Swan, la hipótesis de convergencia beta condicional puede ser testada utilizando la siguiente ecuación (Barro y Sala-i-Martin, 1995):

$$\gamma_{it} = \alpha + \beta \log(y_{i,t-1}) + \varphi' X_{i,t-1} + u_{it} \quad (1)$$

Donde: $y_{i,t-1}$: PBI p.c. del país i ésimo en el período temporal $t-1$.

$\gamma_{it} = \log(y_{it}) - \log(y_{i,t-1})$: tasa de crecimiento del PBI p.c.

$X_{i,t-1}$: vector de variables que determinan el estado estacionario de cada país.

u_{it} : perturbación aleatoria de variables IID $\sim(0, \sigma^2)$.

La muestra seleccionada es tipo datos de panel anuales, en donde las cuentas nacionales fueron tomadas de CEPAL y las variables representativas del estado estacionario, de la HPnStats del Banco Mundial, habiéndose utilizado: tasa de crecimiento de la inversión, tasa de crecimiento de la población, apertura de la economía (exportaciones + importaciones hacia los países del Mercosur sobre exportaciones + importaciones totales), nivel de educación secundaria y otras.

Se realizó una estimación preliminar utilizando paneles con efectos fijos, metodología que toma en cuenta la existencia de heterogeneidades individuales correlacionadas con los regresores (estrictamente exógenos). Las estimaciones citadas en este trabajo corresponden a Brufman, Martínez y Urbisaia (2011); pero no comentaremos los resultados obtenidos sino después de estudiar y discutir la metodología econométrica que nos ocupa.

3.1 Cuestiones econométricas relacionadas con la aplicación de un modelo simple de efectos fijos

Al estimar una ecuación de carácter netamente dinámico en el tiempo, esta metodología no prevé la posible presencia de los siguientes problemas:

1. El problema de la autocorrelación:

- la endógena rezagada: introduce correlación y endogeneidad. Los errores estándares MCO no son generalmente válidos.
- efectos fijos son consistentes sólo bajo exogeneidad estricta.
- es menester estudiar el patrón de autocorrelación potencialmente presente en la ecuación poblacional.
- el dinamismo del problema económico (estimar parámetros del estado estacionario inobservable) es en sí mismo una fuente potencial de autocorrelación.

2. A pesar de estar trabajando con datos de panel, nuestra muestra es potencial candidata a presentar problemas propios de las *series de tiempo*.

3. Establecemos el supuesto de que las heterogeneidades no observables están correlacionadas con las variables explicativas del vector X . No obstante, es un supuesto que estaríamos dispuestos a sostener, dado que los factores socio-económicos de los países de nuestra muestra tienen relación con las variables del vector X y por eso nos permitiríamos removerlos¹.

4. Si existe endogeneidad, y nuestro objetivo es estimar un modelo dinámico, necesitamos respecificar nuestro modelo o encontrar la forma de hacer inferencia aun en presencia de autocorrelación. Frente a estos problemas, las alternativas disponibles serían: construir un modelo de ecuaciones simultáneas, utilizar variables instrumentales o variables proxy. Como no es nuestro objetivo explicar el comportamiento de cada una de las variables involucradas en las estimaciones sino comprobar la existencia o no de convergencia condicional, no tiene sentido desarrollar un modelo de ecuaciones simultáneas. Es en este contexto en el cual las variables instrumentales juegan un papel importante a la hora de solucionar el problema de la endogeneidad. GMM en diferencias y GMM en sistemas son estimadores generalizados de momentos que utilizan variables instrumentales.

3.2 Estimador 1: Paneles dinámicos: El estimador GMM en diferencias de Arellano y Bond

Para aplicar esta metodología, respecificamos la ecuación presentada en (1) de la siguiente forma:

$$y_{it} = \beta y_{i,t-1} + \alpha' X_{it}^* + \eta_i + v_{it} = \gamma' X_{it} + \eta_{it} + v_{it} \quad (2)$$

donde: $X_{it} = [y_{i,t-1}, X_{it}^*]'$; η_{it} es el efecto fijo por unidad de sección transversal y v_{it} es la perturbación aleatoria. Las variables se toman en logaritmos.

(2) puede ser estimada en primeras diferencias con la siguiente reparametrización:

$$\Delta y_{it} = \delta \Delta y_{i,t-1} + \alpha' \Delta X_{it}^* + \Delta v_{it} \quad (3)$$

¹ Nuestra preocupación es meramente formal, debido a posibles inconsistencias del test de Hausman (1978).

En esta especificación, el coeficiente δ corresponde a $(1+\beta)$. De esta manera, GMM se aplica a primeras diferencias y la matriz de instrumentos Z tratará de resolver la endogeneidad introducida por la endógena rezagada y otros regresores no estrictamente exógenos.

La evaluación de los resultados obtenidos incluye el control de:

- a) la posible persistencia de autocorrelación en los residuos.
- b) la validez de los instrumentos seleccionados, con énfasis en las X_{it} .

Teniendo en cuenta la imposibilidad de facto de saber si estamos trabajando con una perturbación esférica y regresores estrictamente exógenos, será altamente deseable el uso de una matriz Z pertinente, de manera tal de satisfacer las condiciones de ortogonalidad entre los regresores y la perturbación aleatoria.

Una cuestión econométrica básica en nuestra ecuación poblacional, es distinguir qué patrón potencial de autocorrelación podría estar presente en la perturbación aleatoria. Para eso, distinguimos las siguientes posibles situaciones:

- si la ecuación dinámica en niveles tuviera una perturbación tipo ruido blanco, su expresión en primeras diferencias tendría estructura $MA(1)$, esto es, un proceso estacionario cuya matriz de varianzas-covarianzas mostraría autocorrelación de primer orden, pero no de segundo orden.
- si la ecuación dinámica en niveles tuviera una perturbación tipo random walk, entonces las primeras diferencias de esta perturbación serían ruido blanco.
- una tercera alternativa es que la perturbación responda a un proceso autorregresivo de primer orden; esta situación es poco probable porque el modelo contiene como variable explicativa la endógena rezagada.

La estrategia desarrollada por el estimador GMM en diferencias consiste en:

- testear la posible existencia de autocorrelación en los residuos de 1° y 2° orden.
- si la matriz de instrumentos Z es apropiada (de manera tal que constituye una solución al problema de la endogeneidad), no existirá autocorrelación de 2do orden o superior en la ecuación en primeras diferencias.
- el test de Sargan-Hansen completa la evaluación global de la estimación.

3.3 Estimador 2: Paneles dinámicos: El estimador GMM en sistemas de Arellano-Bover (1995) y Blundell-Bond (1998)

Tras haber abordado el problema de la autocorrelación introducida por la endógena rezagada y los potenciales regresores no endógenos en el estimador anterior, surge la inquietud ante problemas econométricos remanentes que no han sido cubiertos por GMM en diferencias:

-la posibilidad de que las variables resulten ser no estacionarias.

-la posibilidad de que exista no cointegración de las anteriores, en la ecuación.

-los instrumentos propuestos en GMM en diferencias resultan usualmente débiles si la endógena rezagada tiende a uno y la varianza relativa σ_η^2/σ_u^2 tiende a incrementarse (Blundell y Bond, 1998).

Si planteamos nuestra ecuación (1) como un panel de datos de efectos aleatorios, esta ecuación puede expresarse como:

$$\begin{aligned} y_{it} &= \beta y_{i,t-1} + \alpha_1' x_{it} + \alpha_2' x_{i,t-1} + v_{it} = \\ &= \beta y_{i,t-1} + \alpha_1' x_{it} + \alpha_2' x_{i,t-1} + \eta_i + u_{it} \end{aligned} \quad (4)$$

donde $v_{it} = \eta_i + u_{it}$ representa la clásica descomposición del error en: heterogeneidad no observada η_i , incorrelacionada con las variables explicativas, y la perturbación idiosincrática u_{it} con las condiciones estándares:

$$\eta_i \sim IID(0, \sigma_\eta^2); u_{it} \sim IID(0, \sigma_u^2); E(\eta_i u_{it}) = 0 \text{ para } i = 1, \dots, N \text{ y } t = 2, \dots, T.$$

Se supone además al inicio que:

$$y_{i1} = \frac{\eta_i}{1-\beta} + v_{i1} \quad (5)$$

(5) introduce una condición inicial de estacionariedad necesaria para lograr un nivel de convergencia para y_{it} a partir de $t=2$. El requerimiento clave es que las discrepancias de las condiciones iniciales respecto al nivel $\eta_i/(1-\beta)$ estén incorrelacionadas con dicho nivel.

Frente al problema de endogeneidad de los regresores, los autores sugieren emplear restricciones adicionales de momentos, utilizando dos tipos de variables instrumentales: en nivel para variables diferenciadas y diferenciadas para variables en nivel.

Llegados a este punto, aparece una nueva cuestión: ¿son nuestras series persistentes o no? Para superar este interrogante, aplicamos tests de raíces unitarias para series múltiples en datos de panel, desarrollados por Levin, Lin and Chu (2002), Im, Pesaran y Shin (2003), entre otros, teniendo en cuenta la posible existencia de raíces unitarias individuales o comunes e interceptos individuales o comunes. Aplicadas a las series de nuestra muestra, la mayoría de las series son integradas de orden (1), excepto la tasa de inversión y el nivel de analfabetismo, que resultaron estacionarias.

3.4 Otra cuestión econométrica

¿Es posible estimar una ecuación satisfactoriamente, sin saber si las variables integrantes cointegran? La investigación reciente al respecto nos dice lo siguiente:

- Blundel y Bond (1998): al simular series para su estimador con covarianzas heterocedasticas y ausencia de condición inicial de estacionariedad, concluyeron que el estimador conserva su propiedad de consistencia en esas circunstancias.
- Kruiniger (2007): en una estimación de datos de panel dinámicos, con covarianza estacionaria y en presencia de raíces unitarias, el estimador lineal GMM es súper consistente.
- Baltagi y Kao (2000), Choi (2006) y Breitung y Pesaran (2007): proponen distintos tests de cointegración para datos de panel³.

A continuación, se estimó la ecuación (4) con GMM en sistemas, el cual, en resumen, aborda las siguientes cuestiones: la introducción de autocorrelación en la perturbación aleatoria por los motivos ya mencionados; la existencia de heterogeneidades no observables; la presencia de raíces unitarias en la variable PBI p.c. y en la mayoría de los regresores. Si los tests de autocorrelación y de Sargan-Hansen de validez de la matriz de instrumentos resultan satisfactorios, podemos darnos por satisfechos de que arribamos a un **sistema estable**, debido a que el sistema cointegra.

³ Esta alternativa no fue explorada en este trabajo.

3.5 ¿GMM en diferencias o GMM en Sistemas?

Las estimaciones para nuestro coeficiente β obtenidas de acuerdo a las metodologías mencionadas fueron las siguientes:

	Alternativa A	Alternativa B	Alternativa C
Efectos fijos	-0,03	-0,02	-0,02
GMM en diferencias	-0,08	-0,13	-0,03
GMM en sistemas	+0,04	+0,0009	-0,12

Tabla 1. Resultados (Brufman, *et. al.*, 2011)

Cada alternativa corresponde a diferentes regresores utilizados como representativos de la posición de estado estacionario (variables proxies del desarrollo humano, pero no variando las especificaciones (2) y (4)). Efectos fijos trabaja como una aproximación preliminar a la estimación. A la vista de estos resultados tan disímiles, y antes de inclinarnos por uno en particular, es conveniente plantear una discusión acerca de la bondad de los estimadores GMM en sus dos versiones:

1. GMM en diferencias intenta resolver como problema fundamental en paneles dinámicos la presencia de regresores no estrictamente exógenos. Decidir a priori si los regresores presentan o no esta condición es una tarea difícil, por lo que esta metodología propone analizar los residuos para validar la estimación global. Cuando usamos GMM en diferencias, no necesitamos preocuparnos a priori por la correcta clasificación de las variables involucradas.
2. GMM en diferencias puede producir instrumentos débiles si el parámetro de interés es cercano a uno.
3. Debido a la persistencia de la endógena rezagada, GMM en sistemas resulta más adecuado a nuestro problema económico. Pero aun así, el estimador introduce algunas arbitrariedades en su construcción.
4. Suponer o testear que las heterogeneidades no observables se comportan como efectos fijos o aleatorios influirá en la decisión final.

4. CONCLUSIONES

La aplicación del test de efectos fijos redundantes nos inclina hacia GMM en sistemas, conjuntamente con la condición de las variables involucradas en la muestra, integradas de orden 1 en su mayoría.

¿Hemos encontrado evidencia de convergencia condicional en nuestra muestra? El análisis exploratorio de datos y la evolución histórica del Mercosur nos indican ausencia de convergencia condicional. En las estimaciones GMM en sistemas, sólo la tercera alternativa sugiere no convergencia condicional.

Si nos atenemos a los tests aplicados y ya comentados, las especificaciones estimadas por GMM en diferencias y GMM en sistemas resultan adecuadas desde el punto de vista econométrico, y en especial esta última. Pero, ¿a qué conclusión arribamos con respecto a la existencia o no de convergencia β -condicional en nuestra muestra de países del Mercosur? La estimación del parámetro β no es concluyente. La evidencia sugiere que existe convergencia condicional en el período considerado, pero o bien las tasas de convergencia son muy cercanas a cero, o en determinados tramos de la muestra existieron períodos de no convergencia, resultando en promedio una conclusión ambigua.

Más allá de estas conclusiones netamente empíricas, cabe observar que los intervalos de confianza para los parámetros estimados incluyen siempre al cero como valor más o menos intermedio. Llegados a este punto, cabe preguntarse si las metodologías econométricas utilizadas en este trabajo y en estudios previos de otros autores para el estudio de la convergencia condicional son realmente aplicables en este caso, en el cual nuestra conclusión depende crucialmente del signo negativo o positivo de la estimación efectuada. El problema de la convergencia económica ha motivado gran parte del desarrollo de paneles dinámicos.

Los estimadores GMM en diferencias y GMM en sistemas resultan adecuados para la estimación de ecuaciones dinámicas. Sin embargo, y a pesar de haberse aplicado extensamente al problema de la convergencia económica, ante la sensibilidad de la respuesta al interrogante principal ("existencia o no de convergencia") al signo obtenido, y teniendo en cuenta que las estimaciones obtenidas en este trabajo y en los anteriores son siempre muy cercanas a cero, el modelo econométrico en general no resulta una opción adecuada para este problema empírico en particular. La tarea a futuro de los investigadores es replantear la contrastación empírica de este problema económico con un modelo estadístico o econométrico diferente al que se ha estado utilizando en los '90 y '00.

BIBLIOGRAFÍA

- Anderson, T.W. y Hsiao, C. (1981) "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, vol. 76, pp.598-606.
- Angrist, J.D. y A.B. Krüger (2001) "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments". *Journal of Economic Perspectives* vol. 15, N°4, pp.69-85.
- Arellano, M y S. Bond (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". *Review of Economic Studies*, vol. 58, pp.277-297.
- Arellano, M. y Bover, O. (1995) "Another look at the instrumental variable estimation of error-components models" *Journal of Econometrics*, vol. 68, pp.29-51.
- Baltagi, B.H. (2005) *Econometric Analysis of Panel Data*, 3rd. Edition. Wiley & Sons. Chichester, England.
- Bhargava, A. y Sargan, J.D. (1983): "Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods", *Econometrica*, vol. 51, pp.1635-1659.
- Barro, R. y Sala-i-Martin, X. (1995). *Economic growth*. McGraw-Hill, Estados Unidos.
- Blundell, R y Bond, S. (1998) "Initial conditions and moment restrictions in dynamic panel models". *Journal of Econometrics*, vol. 87, pp.115-143.
- Bond S.; Hoeffler, A; y Temple, J. (2001) "GMM Estimation of Empirical Growth Models". *Centre for Economic Policy Research*. Discussion Paper N°3048.
- Breitung, J. (2000). "The Local Power of Some Unit root Tests for Panel Data", en Baltagi: *Advances in Econometrics*, vol. 15. JAI Press, Amsterdam, pp.161-178.
- Brufman, J.; Martínez, C.; y Urbisaia, H. (2011) "Estimadores GMM en paneles dinámicos. Aplicación al estudio de convergencia beta en el Mercosur." *Documentos de Trabajo de la Sección de Estadística y Econometría*, Facultad de Ciencias Económicas, Universidad de Buenos Aires, N°9.
- CEPAL (2009). *Bases de datos sobre estadísticas varias*. www.eclac.cl
- Drukker, D. (2008). "Econometric analysis of dynamic panel-data models using Stata". *Summer North American Stata Users Group meeting*.

- Durlauf, S.; Johnson, P. y Temple, J. (2004): "Growth Econometrics." *Vassar College Economics Working Paper* N°61.
- Hadri, K. (2000) "Testing for Stationarity in Heterogeneous Panel Data", *Econometric Journal*, vol. 3, pp.148-161.
- Hansen, L. (1982) "Large sample properties of generalized method of moments estimators". *Econometrica*, vol. 50, pp.1029-1054.
- Hausman, J.A. (1978) "Specification tests in Econometrics." *Econometrica* vol. 46, pp.1251-1271.
- Hsiao, C. (2003) *Analysis of Panel Data*. 2nd Edition. Cambridge University Press. Cambridge, U.K.
- IM, K; Pesaran, M. Y Shin, Y. (2003) "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, vol. 115, pp.53-74.
- Levin, A.; Lin, C. y Chu, C. (2002) "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", *Journal of Econometrics*, vol. 108, pp.1-24.
- Nickell, S.J. (1981) "Biases in Dynamic Models with Fixed Effects". *Econometrica*, vol. 49, pp.1417-26.
- Roodman, D. (2006) "How to do xtabond2: An Introduction to "Difference" and "System" GMM in Stata. *Working Paper* N°103, *Center for Global Development*. www.cgdev.org
- World Bank (2009). *Base de datos HPnStats*. www.worldbank.org